

# IMPLICATIONS OF MIXING PROPERTIES ON INTRINSIC ERGODICITY IN SYMBOLIC DYNAMICS

RONNIE PAVLOV

**ABSTRACT.** We study various mixing properties for subshifts, which allow words in the language to be concatenated into a new word in the language given certain gaps between them. All are defined in terms of an auxiliary gap function  $f : \mathbb{N} \rightarrow \mathbb{N}$ , which gives the minimum required gap length as a function of the lengths of the words on either side. In this work we focus mostly on topological transitivity, topological mixing, and a property which we call non-uniform two-sided specification.

It was shown in [12] that non-uniform specification with gap function  $\Omega(\ln n)$  does not imply uniqueness of the measure of maximal entropy. In this work, we show that for all other gap functions, i.e. for all  $f(n)$  with  $\liminf \frac{f(n)}{\ln n} = 0$ , even the weaker property of non-uniform two-sided specification implies uniqueness of the measure of maximal entropy, and its full support. We also show that even the very weak property of topological transitivity implies uniqueness of the measure of maximal entropy (and its full support) when  $\lim \frac{f(n)}{\ln n} = 0$ . We then define a class of subshifts satisfying this hypothesis, each of which therefore has a unique measure of maximal entropy.

We also give classes of examples which demonstrate two negative results. The first is that multiple measures of maximal entropy can occur for subshifts with topological mixing with gap function which grows arbitrarily slowly along a subsequence. The second is that topological mixing can occur with arbitrarily slowly growing unbounded gap function for subshifts without any periodic points.

## 1. INTRODUCTION

In this work, we examine the implications of various “mixing” conditions in symbolic dynamics. In symbolic dynamics, sequences of symbols from a finite alphabet are acted on by shifts, and a set of such sequences which is closed and invariant under shifts is called a subshift. A subshift can be equivalently defined via the set of all finite strings of symbols from the alphabet (or words) which appear within its sequences; the set of such words is called the language of the subshift. Mixing conditions on a subshift all have the same basic structure: given some words in the language, you want to be able to combine them to create a new word in the language, given a gap between them at least as large as some threshold. The main result of this work is that for all such conditions, if the required gap has length which grows sublogarithmically in the length of the words to be combined, then the subshift has a unique measure of maximal entropy.

There are several simple choices one can make in the definition of a mixing condition, all of which yield different properties. A first choice is whether the

---

2010 *Mathematics Subject Classification.* Primary: 37B50; Secondary: 37B10, 37A35.

*Key words and phrases.* Symbolic dynamics, non-uniform specification, intrinsic ergodicity.

The author gratefully acknowledges the support of NSF grant DMS-1500685.

thresholds in question can always be chosen at a uniform length, or whether they may need to increase with the lengths of the words to be combined. We will call these uniform and non-uniform conditions.

Another distinction can be drawn between so-called topological mixing conditions and specification conditions. Topological mixing simply guarantees that one can interpolate between a pair of words, given a gap not smaller than their associated threshold. Specification conditions, on the other hand, guarantee interpolation for an arbitrary collection of words when gaps are placed between each pair which are not smaller than their associated thresholds. This difference is subtle, but important: for a system with only topological mixing, combining more and more  $n$ -letter words may eventually require larger and larger gaps, whereas for a system with specification, the same gap could be used between each pair.

In addition, the thresholds may depend on the lengths of words on both sides, or only on one (say the left without loss of generality); we can refer to these as “two-sided” or “one-sided” conditions respectively.

This would seem, a priori, to yield eight different conditions: (uniform or non-uniform) (one-sided or two-sided) (topological mixing or specification). It is quickly evident that some of the conditions are equivalent though. For instance, in the uniform case, the one-sided/two-sided distinction is irrelevant, since the threshold in fact does not depend on the words to be combined at all. In fact the topological mixing/specification distinction also turns out to be irrelevant in this case. To wit, under the assumption of uniformity, an arbitrary collection of words separated by gaps not smaller than the uniform threshold can be combined one pair at a time; first two adjacent words are combined, then the newly created word is combined with a third, etc. The uniform case really yields only a single condition then.

In the non-uniform case, two conditions still collapse: non-uniform one-sided topological mixing in fact implies non-uniform one-sided specification. The reasoning is the same as above: given a collection of words where each pair is separated by a gap not smaller than the threshold determined by the word on its left, one can first combine the rightmost two, then combine the newly created word with the word immediately to the left (the gap between them still suffices since the word to the left is unchanged from the original collection), and so on.

This leaves four legitimately different conditions: in decreasing order of strength, these are uniform topological mixing/specification, non-uniform one-sided specification, non-uniform two-sided specification, and non-uniform two-sided topological mixing. The first condition was originally defined by Bowen ([1]) in greater generality, and referred to simply as **specification**. The second condition was originally defined by Marcus ([10]) and not given a name; he used it (in a more general setting) to study ergodic toral automorphisms. In other literature, this has been referred to by several names, including almost weak specification ([3]), weak specification ([8]), and non-uniform specification ([12]). In this work, we simply omit the adjective one-sided and refer to this property as **non-uniform specification**. The third condition is, to our knowledge, new, and so we use the new name **non-uniform two-sided specification**. The fourth condition is traditionally referred to only as **topological mixing** (see [9], [13]), and we follow that convention here. In fact we add just one more condition to the list; we will say that a subshift satisfies **topological transitivity** if any two words can be interpolated at a single distance less than or equal to the threshold (rather than at all distances not smaller than the

threshold); clearly topological transitivity is even weaker than topological mixing. (See Section 2 for formal definitions of all of these properties.)

Specification, the strongest of these properties, is known to imply many useful properties (even for more general topological dynamical systems, see [1]), including the uniqueness of the measure of maximal entropy, or **intrinsic ergodicity**. It has been well-known for some time (see [4], [6], [7]) that in general, topological mixing does not imply intrinsic ergodicity. In [2], the question was posed of whether non-uniform specification with gap function  $f(n) = o(n)$  might imply intrinsic ergodicity. This question was answered negatively in both [8] and [12]. In particular, the following was proved in [12].

**Theorem 1.1** ([12], Theorem 1.1). *For any  $f$  with  $\liminf \frac{f(n)}{\ln n} > 0$ , there exists a subshift  $X$  with non-uniform specification with gap function  $f(n)$  and distinct ergodic measures of maximal entropy, with disjoint supports.*

A positive result was also shown there:

**Theorem 1.2** ([12], Theorem 1.3). *If  $X$  is a subshift with non-uniform specification with gap function  $f$  satisfying  $\liminf \frac{f(n)}{\ln n} = 0$ , then  $X$  cannot have distinct ergodic measures of maximal entropy with disjoint supports.*

This shows that there is some sort of a behavioral transition for non-uniform specification when  $f(n)$  is on the order of  $\ln n$ . However, the question was left open of whether there is any unbounded  $f$  at all which guarantees intrinsic ergodicity. Our first main result answers this positively, even for the slightly weaker condition of non-uniform two-sided specification.

**Theorem 1.3.** *If  $X$  has non-uniform two-sided specification with gap function  $f(n)$  with  $\liminf \frac{f(n)}{\ln n} = 0$ , then  $X$  has a unique measure of maximal entropy, which is fully supported.*

In particular, Theorems 1.2 and 1.3 completely answer the question of when non-uniform specification forces uniqueness of the measure of maximal entropy; this is the case if and only if  $\liminf \frac{f(n)}{\ln n} = 0$ .

Topological mixing and transitivity are much weaker conditions, and to our knowledge, it has never been conjectured that either could imply intrinsic ergodicity. We show that under a slightly stronger assumption on  $f$ , in fact this is the case.

**Theorem 1.4.** *If  $X$  has topological transitivity with gap function  $f(n)$  with  $\lim \frac{f(n)}{\ln n} = 0$ , then  $X$  has a unique measure of maximal entropy, which is fully supported.*

The proof techniques for Theorems 1.3 and 1.4 are fairly similar to the proof of Theorem 1.3 from [12]. Specifically, we contradict the existence of measures of maximal entropy  $\mu \neq \nu$  by creating “too many” words in  $\mathcal{L}(X)$  by using a mixing property to combine “large” collections of words based on  $\mu$  and  $\nu$ . In [12], we assumed that the supports of  $\mu$  and  $\nu$  were disjoint, meaning that words from their languages could not overlap above a certain length, which allowed us to ensure that all words created were distinct.

The main idea which replaces the assumption of disjoint supports in this work is the maximal ergodic theorem, which allows us to define “large” collections of words based on  $\mu$  and  $\nu$  so that past a fixed length, no word can be both a prefix of a word from the first collection and a suffix of a word from the second.

We anticipate that Theorem 1.4 can be used to prove intrinsic ergodicity for many subshifts. The following theorem gives a class of such subshifts.

**Theorem 1.5.** *For any alphabet  $A$  with  $|A| > 1$  and any sequence of words  $\{w_n\}$  satisfying*

- (1)  $\frac{\ln |w_n|}{3^n} \rightarrow \infty$ ,
- (2)  $|w_n| > \max(16|w_{n-1}|, 6 \cdot 3^n)$  for  $n > 1$ , and
- (3) no  $w_n$  contains subwords of the form  $w0$  and  $w1$  for  $w$  with length at least  $\frac{|w_n|}{3}$ ,

*the subshift  $X(\{w_n\})$  consisting of all sequences on  $A$  which contain no word in  $\{w_n\}$  has topological mixing with a gap function  $f(n)$  satisfying  $\lim_{n \rightarrow \infty} \frac{f(n)}{\ln n} = 0$ .*

The following corollary is immediate.

**Corollary 1.6.** *Any subshift as in Theorem 1.5 has a unique measure of maximal entropy, which is fully supported.*

**Remark 1.7.** Though we do not claim that the hypotheses in Theorem 1.5 are optimal, we note that some similar assumption(s) are necessary. Firstly, some assumption on the rate of growth of  $|w_n|$  is necessary to even force  $X \neq \emptyset$  (see [11]). Also, no matter how fast  $|w_n|$  is assumed to grow, some assumption besides just fast growth of the lengths  $|w_n|$  is necessary to yield even topological transitivity. For instance, if  $A = \{0, 1\}$ , then for any strictly increasing sequence  $\{m_n\}$ , we can take  $w_1 = 0^{m_1-1}1$ ,  $w_2 = 10^{m_2-1}$ , and  $w_n = 01^{m_n-2}0$  for all  $n > 2$ , and then  $X$  is not topologically transitive. Indeed,  $0^{\mathbb{Z}}, 1^{\mathbb{Z}} \in X$ , and so  $0^{n_2}, 1$  are in the language of  $X$ . However, there is no point of  $X$  containing  $0^{n_2}$  and  $1$  (indeed, the only point of  $X$  containing  $0^{n_2}$  is  $0^{\mathbb{Z}}$ ), precluding topological transitivity of  $X$ .

We also note that the proof of Theorem 1.5 could easily be adapted to allow for a weaker version of (3) where  $\frac{|w_n|}{3}$  is replaced by  $\frac{|w_n|}{2+\epsilon}$  for any  $\epsilon > 0$ , which would then require stronger versions of (1) and (2) with different constants. Since our goal was just to describe a class of examples, we chose for simplicity not to complicate the hypotheses further.

Finally, we show two negative results. The first shows that Theorem 1.4 cannot have hypotheses weakened to anything of the sort in Theorem 1.3 (i.e. which guarantee slow growth only along a subsequence.)

**Theorem 1.8.** *For every increasing unbounded function  $f(n)$ , there exists a function  $g(n)$  with  $\liminf \frac{g(n)}{f(n)} = 0$ , and a subshift  $X$  with topological mixing with gap function  $g(n)$ , zero topological entropy, and multiple invariant measures (which are automatically measures of maximal entropy).*

Under the assumption of the specification property, the unique measure of maximal entropy is in addition known to be a weak limit of averages over periodic orbits. Our second negative result demonstrates that for topological mixing, there is no hope of proving such a fact, since periodic orbits might not exist no matter how slowly  $f(n)$  grows.

**Theorem 1.9.** *For every increasing unbounded function  $f(n)$ , there exists a subshift with topological mixing with gap function  $f(n)$  which contains no periodic points.*

We note that Theorem 1.4 then yields intrinsic ergodicity for some subshifts with no periodic points. We do not know whether non-uniform specification with sufficiently slowly growing  $f$  might force the existence of periodic points.

**Remark 1.10.** We would like to briefly compare and contrast some of our results to the work of Gurevic in [5]. He considered any subshift  $X$  as the decreasing intersection of an associated sequence of shifts of finite type  $\{X_n\}$ ; for any  $n$ ,  $X_n$  is defined as the set of all sequences where each  $n$ -letter subword is in the language of  $X$ . He then considers two sequences of numbers defined by  $\{X_n\}$ . The first,  $\alpha_n$ , is the maximum over pairs of  $n$ -letter words in the language of  $X_n$  of the minimum distance required to interpolate between them in  $X_n$ . Since  $X \subset X_n$ , this is less than or equal to the analogous quantity for pairs of  $n$ -letter words in the language of  $X$ . Also, in his definition, he measures the distance between the initial letters of the two  $n$ -letter words, meaning that his distances are  $n$  units greater than the ones we consider. Summarizing, if  $X$  is topologically transitive with gap function  $f(n)$ , then  $n \leq \alpha_n \leq n + f(n)$ . The second sequence,  $\rho_n$ , is given by the differences  $h(X_n) - h(X)$  between the topological entropies of  $X_n$  and  $X$ . Among other results, it is shown in [5] that if  $h(X) > 0$  and the inequality

$$(1) \quad \rho_n \leq e^{-(16+\epsilon)\alpha_n h(X)}$$

holds for some  $\epsilon > 0$  and infinitely many  $n$ , then  $X$  is intrinsically ergodic.

This result is impressive first of all because it requires information about slow growth of  $\alpha_n$  (relative to  $\rho_n$ ) only along a subsequence. However, the result ignores information about slow growth of a gap function for topological transitivity (in fact,  $f(n) = o(n)$  implies  $\alpha_n = n + o(n)$ , and so for all such  $f$ , exactly the same sequences  $\{\rho_n\}$  satisfy (1) for infinitely many  $n$ . Also, substantial information about  $\rho_n$  is required, whereas it need not be considered for our results.

Examples are also given in [5] of subshifts with  $\alpha_n \leq 3n$  for all  $n$  and multiple measures of maximal entropy, and with  $\alpha_n \leq 2n$  for large enough  $n$  and no periodic points (these examples also illustrate quick decay of  $\rho_n$ , relevant for the results there). Theorems 1.1, 1.8, and 1.9 exhibit the same phenomena with topological mixing rather than just transitivity, and with even lower gap functions (though of course without any information about  $\rho_n$ .)

#### ACKNOWLEDGMENTS

The author would like to thank Jerome Buzzi and Sylvain Crovisier for the useful observation that a previous version of Theorem 1.4, which used the hypothesis of topological mixing, could be easily strengthened to use only topological transitivity.

#### 2. DEFINITIONS

**Definition 2.1.** For any finite alphabet  $A$ , the **full shift** over  $A$  is the set  $A^{\mathbb{Z}} = \{\dots x_{-1}x_0x_1\dots : x_i \in A\}$ , which is viewed as a compact topological space with the (discrete) product topology.

**Definition 2.2.** The **shift action**, denoted by  $\{\sigma^n\}_{n \in \mathbb{Z}}$ , is the  $\mathbb{Z}$ -action on a full shift  $A^{\mathbb{Z}}$  defined by  $(\sigma^n x)_m = x_{m+n}$  for  $m, n \in \mathbb{Z}$ .

**Definition 2.3.** A point  $x \in A^{\mathbb{Z}}$  is **periodic** if there exists  $n > 0$  for which  $\sigma^n x = x$ .

**Definition 2.4.** A **subshift** is a closed subset of a full shift  $A^{\mathbb{Z}}$  which is invariant under the shift action and which is compact with respect to the induced topology from  $A^{\mathbb{Z}}$ .

The single shift  $\sigma := \sigma^1$  is an automorphism on any subshift, and so for any subshift  $X$ ,  $(X, \sigma)$  is a topological dynamical system.

**Definition 2.5.** A **word** over  $A$  is a member of  $A^n$  for some  $n$ , whose **length**  $n$  is denoted by  $|w|$ .

**Definition 2.6.** The **language** of a subshift  $X$ , denoted by  $\mathcal{L}(X)$ , is the set of all words which appear in points of  $X$ . For any  $n \in \mathbb{Z}$ ,  $\mathcal{L}_n(X) := \mathcal{L}(X) \cap A^n$ , the set of words in the language of  $X$  with length  $n$ .

**Definition 2.7.** The **topological entropy** of a subshift  $X$  is

$$h(X) := \lim_{n \rightarrow \infty} \frac{1}{n} \ln |\mathcal{L}_n(X)|.$$

We also need some definitions from measure-theoretic dynamics. All measures considered in this paper will be  $\sigma$ -invariant Borel probability measures on a full shift  $A^{\mathbb{Z}}$ , and we denote the space of such measures by  $\mathcal{M}(X)$ .

**Definition 2.8.** A measure  $\mu$  on  $A^{\mathbb{Z}}$  is **ergodic** if any measurable set  $C$  which is shift-invariant, meaning  $\mu(C \triangle \sigma C) = 0$ , has measure 0 or 1.

Not all  $\sigma$ -invariant measures are ergodic, but a well-known result called the ergodic decomposition shows that any non-ergodic measure can be written as a “weighted average” (formally, an integral) of ergodic measures. Also, whenever ergodic measures  $\mu$  and  $\nu$  are unequal, in fact they must be mutually singular ( $\mu \perp \nu$ ), i.e. there must exist a set  $R$  with  $\mu(R) = \nu(R^c) = 0$ . (See Chapter 6 of [13] for proofs and more information.)

A measure  $\mu$  being ergodic means that for  $f \in L^1(\mu)$ , the so-called ergodic averages  $\frac{1}{N} \sum_{i=0}^{N-1} f(\sigma^i x)$  converge  $\mu$ -a.e. to the “correct” value  $\int f d\mu$ ; this is essentially the content of Birkhoff’s ergodic theorem. We will need the following related result, which deals with the supremum of such averages rather than their limit.

**Theorem 2.9** ([14], Theorem 2 (Maximal Ergodic Theorem)). *For  $f \in L^1(\mu)$ , define  $M^+ f := \sup_{N \in \mathbb{N}} \frac{1}{N} \sum_{i=0}^{N-1} f(\sigma^i x)$ . Then for any  $\lambda \in \mathbb{R}$ ,*

$$\lambda \mu(\{M^+ f(x) > \lambda\}) \leq \int_{M^+ f(x) > \lambda} f d\mu.$$

The following corollary is immediate.

**Corollary 2.10.** *For nonnegative  $f \in L^1(\mu)$  and  $Mf$  as in Theorem 2.9, and any  $\lambda \in \mathbb{R}$ ,*

$$\mu(\{M^+ f(x) \leq \lambda\}) \geq 1 - \frac{\|f\|_1}{\lambda}.$$

We note that by considering the inverse shift  $\sigma^{-1}$  instead, both of these results also hold when  $M^+ f$  is replaced by  $M^- f := \sup_{N \in \mathbb{N}} \frac{1}{N} \sum_{i=0}^{N-1} f(\sigma^{-i} x)$ .

We also need the concept of measure-theoretic entropy.

**Definition 2.11.** For any subshift and word  $w \in \mathcal{L}_n(X)$ , the **cylinder set**  $[w]$  is the set of all  $x \in X$  with  $x_1 x_2 \dots x_n = w$ .

**Definition 2.12.** For any  $\sigma$ -invariant measure  $\mu$  on a full shift  $A^{\mathbb{Z}}$ , the **measure-theoretic entropy** of  $\mu$  is

$$h(\mu) := \lim_{n \rightarrow \infty} \frac{-1}{n} \sum_{w \in A^n} \mu([w]) \ln \mu([w]),$$

where terms with  $\mu([w]) = 0$  are omitted from the sum.

In Definitions 2.7 and 2.12, a standard subadditivity argument shows that the limits can be replaced by infimums; i.e. for any  $n$ ,

$$(2) \quad h(X) \leq \frac{1}{n} \ln |\mathcal{L}_n(X)| \text{ and } h(\mu) \leq \frac{-1}{n} \sum_{w \in A^n} \mu([w]) \ln \mu([w]).$$

The classical Variational Principle (Theorem 8.6, [13]) states that for any topological dynamical system, the topological entropy  $h(X)$  is the supremum of  $h(\mu)$  over  $\mu \in \mathcal{M}(X)$ . The following definition is then natural.

**Definition 2.13.** For any subshift  $X$ , a **measure of maximal entropy** on  $X$  is a measure  $\mu$  with support contained in  $X$  for which  $h(\mu) = h(X)$ .

Every subshift has at least one measure of maximal entropy; the entropy map  $\mu \mapsto h(\mu)$  is upper semi-continuous in that case ([13], Theorem 8.2), and so must achieve its supremum  $h(X)$  on the compact space  $\mathcal{M}(X)$  (endowed with the weak-\* topology). In fact, the ergodic decomposition, along with the fact that the entropy map is affine ([13], Theorem 8.1), implies that the extreme points of the simplex of measures of maximal entropy are precisely the ergodic measures of maximal entropy. In particular, any subshift with multiple measures of maximal entropy in fact has multiple ergodic measures of maximal entropy.

We now move to the mixing properties which we will consider in this work. The following property was originally defined in a more general setting in [10], but was not there given a name.

**Definition 2.14.** A subshift  $X$  has **non-uniform specification with gap function**  $f(n)$  if

- $f(n)$  is nondecreasing
- For any words  $w^{(1)}, w^{(2)}, \dots, w^{(k)} \in \mathcal{L}(X)$ , and for any integers  $n_1, \dots, n_{k-1}$  where  $n_i \geq f(|w^{(i)}|)$  for  $1 \leq i \leq k$ , there exist words  $v^{(1)} \in \mathcal{L}_{n_1}(X), \dots, v^{(k-1)} \in \mathcal{L}_{n_{k-1}}(X)$  so that the word  $w^{(1)}v^{(1)}w^{(2)}v^{(2)} \dots w^{(k-1)}v^{(k-1)}w^{(k)} \in \mathcal{L}(X)$ .

**Definition 2.15.** A subshift  $X$  has **non-uniform two-sided specification with gap function**  $f(n)$  if

- $f(n)$  is nondecreasing
- For any words  $w^{(1)}, w^{(2)}, \dots, w^{(k)} \in \mathcal{L}(X)$ , and for any integers  $n_1, \dots, n_{k-1}$  where  $n_i \geq \max(f(|w^{(i)}|), f(|w^{(i+1)}|))$  for  $1 \leq i < k$ , there exist words  $v^{(1)} \in \mathcal{L}_{n_1}(X), \dots, v^{(k-1)} \in \mathcal{L}_{n_{k-1}}(X)$  so that the word  $w^{(1)}v^{(1)}w^{(2)}v^{(2)} \dots w^{(k-1)}v^{(k-1)}w^{(k)} \in \mathcal{L}(X)$ .

Finally, we consider the simpler properties of topological mixing and topological transitivity.

**Definition 2.16.** A subshift  $X$  has **topological mixing with gap function**  $f(n)$  if

- $f(n)$  is nondecreasing
- For any words  $w, v \in \mathcal{L}(X)$ , and for every  $n \geq \max(f(|w|), f(|v|))$ , there exists a word  $u \in \mathcal{L}_n(X)$  for which  $wuv \in \mathcal{L}(X)$ .

**Definition 2.17.** A subshift  $X$  has **topological transitivity with gap function**  $f(n)$  if

- $f(n)$  is nondecreasing
- For any words  $w, v \in \mathcal{L}(X)$ , there exist  $n \leq \max(f(|w|), f(|v|))$  and a word  $u \in \mathcal{L}_n(X)$  for which  $wuv \in \mathcal{L}(X)$ .

Usually, when topological mixing and transitivity are defined, the dependence of the required gap on the words to be concatenated is not explicitly described in terms of a function, but our results require such a description.

As mentioned in the introduction, non-uniform specification with gap function  $f(n) \Rightarrow$  non-uniform two-sided specification with gap function  $f(n) \Rightarrow$  topological mixing with gap function  $f(n) \Rightarrow$  topological transitivity with gap function  $f(n)$ .

**Remark 2.18.** For these properties, the assumption that  $f(n)$  is nondecreasing is often not explicitly required in the literature. However, it is fairly natural and simplifies our proofs substantially.

### 3. MAIN PROOFS (PROOFS OF THEOREMS 1.3 AND 1.4)

The first tools that we'll need are some upper bounds on the size of  $|\mathcal{L}_n(X)|$  for subshifts with non-uniform two-sided specification or topological transitivity, which generalize the well-known bound  $|\mathcal{L}_n(X)| < Ce^{nh(X)}$  for shifts with specification. In the case of non-uniform two-sided specification, our bound holds for all  $f$ , whereas for topological transitivity, we need to assume that  $f$  grows slowly enough.

**Theorem 3.1.** *If  $X$  is a subshift with non-uniform two-sided specification with gap function  $f(n)$ , then for all  $n$ ,  $|\mathcal{L}_n(X)| \leq e^{(n+f(n))h(X)}$ .*

*Proof.* Suppose that  $X$  is such a shift, and fix any  $n$ . Then, for every  $k$  and any choice  $w_1, \dots, w_k \in \mathcal{L}_n(X)$ , we can use non-uniform two-sided specification to construct a word of the form  $w_1v_1w_2v_2\dots v_kw_k$ , where each  $v_k$  has length  $f(n)$ . Therefore,

$$|\mathcal{L}_{k(n+f(n))}(X)| \geq |\mathcal{L}_n(X)|^k.$$

Taking logarithms of both sides, dividing by  $k(n+f(n))$ , and letting  $k \rightarrow \infty$  yields

$$h(X) \geq \frac{\ln |\mathcal{L}_n(X)|}{n+f(n)},$$

and solving for  $|\mathcal{L}_n(X)|$  completes the proof.  $\square$

**Theorem 3.2.** *If  $X$  is a subshift with topological transitivity with gap function  $f(n)$ ,  $C > 0$ , and  $M \geq 3$  is such that  $f(n) \leq \min(n, C \ln n)$  for all  $n \geq M$ , then for all  $n \geq M$ ,  $|\mathcal{L}_n(X)| \leq 2C(\ln n)e^{(n+C \ln n+C \ln 3)h(X)}$ .*



*Proof.* Suppose that  $X, C, M$  satisfy the hypotheses and fix any  $n \geq M$ . We give lower bounds on  $\mathcal{L}_{n_k}(X)$  for a recursively defined sequence  $n_k$  by using topological transitivity of  $X$ . Define  $n_0 = n$ . Then, for every  $k \geq 0$ , every pair of words in  $\mathcal{L}_{n_k}(X)$  can be combined at some gap less than or equal to  $f(n_k)$ ; this means that there is a single gap length  $i_k$  so that there is a set of such pairs of cardinality at least  $\frac{|\mathcal{L}_{n_k}(X)|^2}{f(n_k)}$  which can all be combined at gap  $i_k$ . Define  $n_{k+1} = 2n_k + i_k$ .

Clearly all  $n_k$  are at least  $M$ , and so by the assumption that  $f(n) \leq n$  for  $n \geq M$ ,  $2n_k \leq n_{k+1} \leq 3n_k$  for all  $n$ . Therefore,  $2^k n \leq n_k \leq 3^k n$  for all  $k$ .

It is easy to prove by induction that for every  $k$ ,  $n_k \leq 2^k n + 2^{k-1} f(n_0) + 2^{k-2} f(n_1) + \dots + f(n_{k-1})$ , and so by monotonicity of  $f$ , for every  $k$

$$n_k \leq 2^k n + 2^{k-1} f(n) + 2^{k-2} f(3n) + \dots + f(3^{k-1} n).$$

Since  $f(n) \leq C \ln n$  for  $n \geq M$ , in fact

$$\begin{aligned} (3) \quad n_k &\leq 2^k n + 2^{k-1} C \ln n + 2^{k-2} C \ln(3n) + \dots + C \ln(3^{k-1} n) \\ &= 2^k \left( n + C \ln n \left( \frac{1}{2} + \frac{1}{4} + \dots + \frac{1}{2^k} \right) + C \ln 3 \left( \frac{1}{4} + \frac{2}{8} + \dots + \frac{k-1}{2^k} \right) \right) \\ &\leq 2^k (n + C \ln n + C \ln 3). \end{aligned}$$

By definition of  $i_k$ ,  $|\mathcal{L}_{n_{k+1}}(X)| \geq \frac{|\mathcal{L}_{n_k}(X)|^2}{f(n_k)}$  for every  $k$ . By induction, for every  $k$ ,  $|\mathcal{L}_{n_k}(X)| \geq |\mathcal{L}_n(X)|^{2^k} \prod_{i=0}^{k-1} f(n_i)^{-2^{k-i-1}}$ , and so

$$\begin{aligned} (4) \quad \frac{\ln |\mathcal{L}_{n_k}(X)|}{n_k} &\geq \frac{2^k \ln |\mathcal{L}_n(X)| - \sum_{i=0}^{k-1} 2^{k-i-1} \ln f(n_i)}{2^k (n + C \ln n + \ln 3)} \\ &\geq \frac{\ln |\mathcal{L}_n(X)| - \sum_{i=0}^{k-1} 2^{-i-1} \ln f(n_i)}{n + C \ln n + C \ln 3}. \end{aligned}$$

We bound the sum somewhat similarly to above, recalling that  $n \geq M \geq 3$ .

$$\begin{aligned} \sum_{i=0}^{k-1} 2^{-i-1} \ln f(n_i) &\leq \frac{1}{2} \ln(C \ln n) + \frac{1}{4} \ln(C \ln(3n)) + \dots + \frac{1}{2^{k-1}} \ln(C \ln(3^{k-1} n)) \\ &\leq \frac{1}{2} \ln(C \ln n) + \frac{1}{4} \ln(C \ln(n^2)) + \dots + \frac{1}{2^{k-1}} \ln(C \ln(n^{2^{k-1}})) = \ln(C \ln n) + \ln 2. \end{aligned}$$

Plugging this into (4) and letting  $k \rightarrow \infty$  yields

$$h(X) \geq \frac{\ln |\mathcal{L}_n(X)| - (\ln(C \ln n) + \ln 2)}{n + C \ln n + C \ln 3},$$

which, upon solving for  $|\mathcal{L}_n(X)|$ , completes the proof.  $\square$

We will also need the following general lower bound on the cardinality of a set of words with positive measure for some measure of maximal entropy.

**Theorem 3.3.** *For any subshift  $X$ , any measure of maximal entropy  $\mu$  on  $X$ , and any  $S \subseteq \mathcal{L}_n(X)$ , if  $|\mathcal{L}_n(X)| \leq e^{h(X)(n+g(n))}$ , then  $|S| \geq e^{nh(X) - \frac{1-\mu(S)}{\mu(S)} g(n)h(X) - \frac{\ln 2}{\mu(S)}}$ .*

*Proof.* Consider such  $X$ ,  $\mu$ , and  $S$ . By (2),

$$(5) \quad nh(X) = nh(\mu) \leq \sum_{w \in A^n} (-\mu(w) \ln \mu(w)) = \sum_{w \in S} (-\mu(w) \ln \mu(w)) + \sum_{w \notin S} (-\mu(w) \ln \mu(w)).$$

For positive  $\alpha_1, \dots, \alpha_N$  with fixed sum  $S$ , it's easily checked that  $\sum(-\alpha_i \ln \alpha_i)$  is maximized when all  $\alpha_i$  are equal (to  $\frac{S}{N}$ ), and so its maximum value is  $S \ln \frac{N}{S}$ . Therefore,

$$\begin{aligned}
 (6) \quad & \sum_{w \in S} (-\mu(w) \ln \mu(w)) + \sum_{w \notin S} (-\mu(w) \ln \mu(w)) \\
 & \leq \mu(S) \ln \left( \frac{|S|}{\mu(S)} \right) + (1 - \mu(S)) \ln \left( \frac{|\mathcal{L}_n(X)| - |S|}{1 - \mu(S)} \right) \\
 & \leq \mu(S) \ln |S| - \mu(S) \ln \mu(S) + (1 - \mu(S)) \ln |\mathcal{L}_n(X)| - (1 - \mu(S)) \ln(1 - \mu(S)) \\
 & \leq \mu(S) \ln |S| + (1 - \mu(S)) \ln |\mathcal{L}_n(X)| + \ln 2 \leq \mu(S) \ln |S| + (1 - \mu(S))(n + g(n))h(X) + \ln 2.
 \end{aligned}$$

(The final inequality uses the assumed upper bound on  $|\mathcal{L}_n(X)|$ .) Combining (5) and (6) yields

$$nh(X) \leq \mu(S) \ln |S| + (1 - \mu(S))(n + g(n))h(X) + \ln 2.$$

Solving for  $\ln |S|$  yields

$$\ln |S| \geq nh(X) - \frac{1 - \mu(S)}{\mu(S)} g(n)h(X) - \frac{\ln 2}{\mu(S)},$$

which upon exponentiation completes the proof.  $\square$

We may now prove our main results.

*Proof of Theorem 1.3.* Suppose for a contradiction that  $X$  has non-uniform two-sided specification for such  $f$  and that  $X$  has multiple measures of maximal entropy. Then, as noted in the introduction,  $X$  has ergodic measures of maximal entropy  $\mu \neq \nu$ , and  $\mu \perp \nu$ . Then, there exists  $R \subset X$  with  $\mu(R) = \nu(R^c) = 0$ . Since  $\mu, \nu$  are Borel measures, there exists  $m$  and  $T$  a union of cylinder sets of words in  $\mathcal{L}_m(X)$  so that  $\mu(T), \nu(T^c) < \frac{1}{5}$ .

Define  $V \subset X$  to be the set of  $x \in X$  for which

$$M^+(\chi_T) = \sup_N \frac{1}{N} \sum_{n=0}^{N-1} \chi_T(\sigma^n x) \leq 2\mu(T).$$

By Corollary 2.10 (to the Maximal Ergodic Theorem),  $\mu(V) \geq 1 - \frac{\int \chi_T d\mu}{2\mu(T)} = \frac{1}{2}$ .

For every  $n$ , define  $V_n$  to be the set of  $n$ -letter words whose cylinder sets have nonempty intersection with  $V$ . For every  $n$ ,  $V_n \supseteq V$ , and so  $\mu(V_n) \geq \frac{1}{2}$ . Also, by definition of  $V$ , for any word  $w$  in some  $V_n$  and any prefix  $p$  of  $w$  with length  $\ell$ , it is the case that  $p$  contains not more than  $2\mu(T)\ell < \frac{2}{5}\ell$  words from  $T$ .

We also note that for every  $n$ , by Theorem 3.3 it is the case that

$$(7) \quad |V_n| \geq e^{h(X)n - h(X)f(n) - 2\ln 2}.$$

Similarly, we define  $W \subset X$  to be the set of  $x \in X$  for which

$$M^-(\chi_{\sigma^m T^c}) = \sup_N \frac{1}{N} \sum_{n=0}^{N-1} \chi_{\sigma^m T^c}(\sigma^{-n} x) \leq 2\nu(T^c)$$

and, for every  $n$ , define  $W_n$  to be the set of  $n$ -letter words whose cylinder sets, after shifting to the left by  $n$  units, have nonempty intersection with  $W$ . By Corollary 2.10 (applied to  $\sigma^{-1}$ ),  $\mu(W) \geq 1 - \frac{\int \chi_{\sigma^m T^c} d\nu}{2\nu(T^c)} = \frac{1}{2}$ , and so for every  $n$ ,

$\nu(W_n) \geq \frac{1}{2}$ . By definition of  $W$ , for any word  $w$  in some  $W_n$  and any suffix  $s$  of  $w$  with length  $\ell$ , it is the case that  $s$  contains not more than  $2\nu(T^c)\ell < \frac{2}{5}\ell$  words from  $T^c$ . As above, for every  $n$

$$(8) \quad |W_n| \geq e^{h(X)n - h(X)f(n) - 2\ln 2}.$$

We note for future reference that no word of length  $\ell \geq 5m$  can possibly be both a prefix of a word in some  $V_n$  and a suffix of a word in some  $W_{n'}$ ; such a word would contain fewer than  $\frac{2}{5}\ell$  words from  $T$  and fewer than  $\frac{2}{5}\ell$  words from  $T^c$ , but its total number of  $m$ -letter subwords is  $\ell - m + 1 > \frac{4}{5}\ell$ , a contradiction.

Now we will complete the proof by using non-uniform two-sided specification of  $X$  to combine words from various  $V_i$  and  $W_j$ , creating so many words in  $\mathcal{L}_n(X)$  that we contradict Theorem 3.1. Choose any  $n$ , and create a family of words in  $\mathcal{L}_n(X)$  as follows. For every integer  $j$  in  $[1, \frac{n-f(n)}{5m+f(n)}]$ , any  $w \in W_{j(5m+f(n))}$ , and any  $v \in V_{n-f(n)-j(5m+f(n))}$ , we use non-uniform two-sided specification to find a word  $u$  of length  $f(n)$  so that  $wuv \in \mathcal{L}_n(X)$ . We make the notation  $g(j, v, w) = wuv$ , and claim that the map  $g$  is injective. To see this, choose any triples  $(j, v, w) \neq (j', v', w')$ . We break into the cases  $j = j'$  and  $j \neq j'$ .

If  $j = j'$ , then either  $v \neq v'$  or  $w \neq w'$ . But  $g(j, v, w) = wuv$  and  $g(j', v', w') = w'u'v'$ , where  $|w| = |w'|$ ,  $|u| = |u'|$ , and  $|v| = |v'|$ . Since  $v \neq v'$  or  $w \neq w'$ , it must be the case that  $g(j, v, w) \neq g(j', v', w')$ .

Now suppose that  $j \neq j'$ , and without loss of generality we assume  $j > j'$ . Then  $g(j, v, w) = wuv$  and  $g(j', v', w') = w'u'v'$ ; suppose for a contradiction that they are equal. Consider their common subword  $z$  occupying the indices from  $j'(5m+f(n)) + f(n) + 1$  through  $j(5m+f(n))$ . Then  $z$  has length  $j(5m+f(n)) - j'(5m+f(n)) - f(n) \geq 5m$ , and is a suffix of  $w$  and a prefix of  $v'$ , a contradiction to the fact noted above. Therefore,  $g(j, v, w) \neq g(j', v', w')$  in this case as well, and  $g$  is injective.

Since  $g$  is an injection to  $\mathcal{L}_n(X)$ ,

$$|\mathcal{L}_n(X)| \geq \sum_{j=1}^{\lfloor (n-f(n))/(5m+f(n)) \rfloor} |W_{j(5m+f(n))}| |V_{n-f(n)-j(5m+f(n))}|.$$

Combining with (7) and (8) yields

$$\begin{aligned} |\mathcal{L}_n(X)| &\geq \left\lfloor \frac{n-f(n)}{5m+f(n)} \right\rfloor (e^{h(X)j(5m+f(n)) - h(X)f(j(5m+f(n))) - 2\ln 2}) \\ &\quad (e^{h(X)(n-f(n)-j(5m+f(n))) - h(X)f(n-f(n)-j(5m+f(n))) - 2\ln 2}) \\ &\geq \left\lfloor \frac{n-f(n)}{5m+f(n)} \right\rfloor e^{nh(X) - 3h(X)f(n) - 4\ln 2}. \end{aligned}$$

(In the final inequality, we used the fact that  $f(j(5m+f(n))), f(n-f(n)-j(5m+f(n))) \leq f(n)$  due to monotonicity of  $f$ .) Then, by Theorem 3.1,

$$e^{(n+f(n))h(X)} \geq \left\lfloor \frac{n-f(n)}{5m+f(n)} \right\rfloor e^{nh(X) - 3h(X)f(n) - 4\ln 2}.$$

We rewrite as

$$(9) \quad e^{4h(X)f(n) + 4\ln 2} \geq \left\lfloor \frac{n-f(n)}{5m+f(n)} \right\rfloor$$

for every  $n$ . Recall, however, that  $\liminf \frac{f(n)}{\ln n} = 0$ , and so there exists a sequence  $n_k$  for which  $\frac{f(n_k)}{\ln n_k} \rightarrow 0$ . Then the left-hand side of (9) behaves as  $n_k^{o(1)}$  and the right behaves as  $n_k^{O(1)}$ , which clearly contradicts (9) for large enough  $k$ . Therefore, our original assumption of multiple measures of maximal entropy on  $X$  was false, and  $X$  has a unique measure of maximal entropy, let's call it  $\mu$ .

It remains to show that  $\mu$  is fully supported, and so for a contradiction assume that there is a word  $y \in \mathcal{L}(X)$  with  $\mu([y]) = 0$ , denote its length by  $\ell$ . Then if we define the subshift  $X_y$  consisting of all points of  $X$  with no occurrences of  $y$ , obviously  $\mu$  is a measure on  $X_y$ , and so  $h(X_y) \geq h(\mu) = h(X)$ . Since  $X_y \subseteq X$ , this means that  $h(X_y) = h(X)$ . Then by (2), for all  $n$

$$(10) \quad |\mathcal{L}_n(X_y)| \geq e^{nh(X)}.$$

Now, we will again use non-uniform two-sided specification to obtain a contradiction to Theorem 3.1. Choose any  $n$ , and create a family of words in  $\mathcal{L}_n(X)$  as follows. For every integer  $j$  in  $[1, \frac{n-2f(n)-\ell}{f(n)+\ell}]$ , any  $w \in \mathcal{L}_{j(f(n)+\ell)}(X_y)$ , and any  $v \in \mathcal{L}_{n-2f(n)-\ell-j(f(n)+\ell)}(X_y)$ , we use non-uniform two-sided specification to find words  $t, u$  of length  $f(n)$  so that  $wtyuv \in \mathcal{L}_n(X)$ . We make the notation  $h(j, v, w) = wtyuv$ , and claim that the map  $h$  is injective. To see this, choose any triples  $(j, v, w) \neq (j', v', w')$ . If  $j = j'$ , then  $h(j, v, w) \neq h(j', v', w')$  exactly as shown above for  $g$ .

If  $j \neq j'$ , then without loss of generality we assume  $j > j'$ . Then  $h(j, v, w) = wtyuv$  and  $h(j', v', w') = w't'yu'v'$ ; suppose for a contradiction that they are equal. Consider their common subword occupying the indices from  $j'(f(n)+\ell) + f(n) + 1$  through  $j(f(n)+\ell) + f(n) + \ell$ . From its location in  $wtyuv$ , this word is  $y$ . However, from its location in  $w't'yu'v'$ , it is a subword of  $v'$ , which was assumed to be in  $\mathcal{L}(X_y)$ , and so we have a contradiction. Therefore,  $h(j, v, w) \neq h(j', v', w')$  in this case as well, and  $g$  is injective.

Since  $g$  is an injection, it yields

$$|\mathcal{L}_n(X)| \geq \sum_{j=1}^{\lfloor \frac{n-2f(n)-\ell}{f(n)+\ell} \rfloor} |\mathcal{L}_{j(f(n)+\ell)}| |\mathcal{L}_{n-2f(n)-\ell-j(f(n)+\ell)}(X_y)|.$$

Using (10) yields

$$(11) \quad |\mathcal{L}_n(X)| \geq \left\lfloor \frac{n-2f(n)-\ell}{f(n)+\ell} \right\rfloor e^{h(X)(n-2f(n)-\ell)}.$$

However, exactly as above, this contradicts Theorem 3.1 for large enough  $k$  if  $\frac{f(n_k)}{\ln n_k} \rightarrow 0$ . Therefore, our assumption was incorrect, and  $\mu$  is fully supported, completing the proof.  $\square$

The proof of Theorem 1.4 is very similar, but requires some slightly different bounds, coming from Theorem 3.2 rather than Theorem 3.1.

*Proof of Theorem 1.4.* Choose  $X$  with topological transitivity for such  $f$ . Take any constant  $C > 0$  (whose value we will fix later.) Since  $\lim \frac{f(n)}{\ln n} = 0$ , we can choose  $M$  so that  $f(n) \leq \min(C \ln n, n)$  for  $n \geq M$ . Assume for a contradiction that  $X$  has multiple measures of maximal entropy; as before, this allows us to define ergodic

measures of maximal entropy  $\mu \perp \nu$ . Define the sets  $V_n$  and  $W_n$  exactly as in the proof of Theorem 1.3. Theorems 3.2 and 3.3 yield the bounds

$$(12) \quad |V_n|, |W_n| \geq (e^{h(X)n-h(X)(C \ln n + C \ln 3) - 2 \ln 2}) / (2C \ln n)$$

for all  $n \geq M$ .

We then proceed similarly as in the proof of Theorem 1.3, creating a family of words in  $\mathcal{L}_n(X)$  via topological transitivity. Specifically, choose any  $n$  for which  $n > 2M + f(n)$ , and for every integer  $j$  in  $[\frac{M}{5m+f(n)}, \frac{n-M-f(n)}{5m+f(n)}]$ , any  $w \in W_{j(5m+f(n))}$ , and any  $v \in V_{n-f(n)-j(5m+f(n))}$ , we use topological transitivity of  $X$  to find a word  $u$  of length  $i_{v,w} \leq \max(f(j(5m+f(n))), f(n-f(n)-j(5m+f(n)))) \leq f(n)$  so that  $wuv \in \mathcal{L}(X)$ . Then, simply extend on the right to a word  $wuvt \in \mathcal{L}_n(X)$ , and denote  $g(j, v, w) = wuvt$ . At this point, we restrict the domain; choose the gap  $i_{v,w} \leq f(n)$  which was used for the maximal number of pairs, and remove all other pairs from the domain of  $g$ . We claim that  $g$  is an injection on the restricted domain, and to that end assume that  $(j, v, w) \neq (j', v', w')$ . The case  $j = j'$  yields  $g(j, v, w) \neq g(j', v', w')$  exactly as in Theorem 1.3 since  $i_{v,w} = i_{v',w'}$  after restriction of the domain. When  $j \neq j'$ , assume without loss of generality that  $j > j'$ . If  $g(j, v, w) = g(j', v', w')$ , a contradiction is reached as before by considering the subword occupying indices from  $j'(5m+f(n)) + i_{v',w'} + 1$  through  $j(5m+f(n))$ , which would be both a suffix of  $w$  and a prefix of  $v'$ . Therefore,

$$|\mathcal{L}_n(X)| \geq \sum_{\lceil M/(5m+f(n)) \rceil}^{\lfloor (n-M-f(n))/(5m+f(n)) \rfloor} \frac{1}{f(n)} |W_{j(5m+f(n))}| |V_{n-f(n)-j(5m+f(n))}|.$$

Since  $j(5m+f(n)), n-f(n)-j(5m+f(n)) \geq M$  for all  $j$ , we can use (12), yielding

$$\begin{aligned} |\mathcal{L}_n(X)| &\geq \left\lfloor \frac{n-2M-f(n)}{5m+f(n)} \right\rfloor (e^{h(X)j(5m+f(n))-h(X)(C \ln(j(5m+f(n))+C \ln 3)-2 \ln 2)} \\ &\quad (e^{h(X)(n-f(n)-j(5m+f(n)))-h(X)(C \ln(n-f(n)-j(5m+f(n))+C \ln 3)-2 \ln 2)} / (f(n)(2C \ln n)^2) \\ &\geq \left\lfloor \frac{n-2M-f(n)}{f(n)(5m+f(n))} \right\rfloor e^{nh(X)-h(X)(3C \ln n+2C \ln 3)-4 \ln 2} / (2C \ln n)^2. \end{aligned}$$

(Here, the last inequality used the fact that  $f(n) \leq C \ln n$  for  $n \geq M$ .) Combining with Theorem 3.2 yields

$$\begin{aligned} &2C(\ln n)e^{(n+C \ln n+C \ln 3)h(X)} \\ &\geq \left\lfloor \frac{n-2M-f(n)}{f(n)(5m+f(n))(2C \ln n)^2} \right\rfloor e^{nh(X)-h(X)(3C \ln n+2C \ln 3)-4 \ln 2} / (2C \ln n)^2. \end{aligned}$$

We rewrite as

$$(13) \quad (2C \ln n)^3 e^{4h(X)C \ln n+3h(X)C \ln 3+4 \ln 2} \geq \left\lfloor \frac{n-2M-f(n)}{f(n)(5m+f(n))} \right\rfloor$$

for every sufficiently large  $n$  (dependent on  $C$ ). If  $h(X) = 0$ , then (13) is a contradiction for large  $n$ , and so we know  $h(X) > 0$ . Then we may choose some  $C < \frac{1}{4h(X)}$ , and note that the left-hand side of (13) behaves as  $n^{4h(X)C+o(1)}$ , while the right behaves as  $n^{O(1)}$ , which yields a contradiction for large enough  $n$ . Therefore, our original assumption of multiple measures of maximal entropy on  $X$  was false, and so  $X$  has a unique measure of maximal entropy  $\mu$ .

It remains to show that  $\mu$  is fully supported, which will be done in exactly the same way as in Theorem 1.3. Assume for a contradiction that there is a word  $y \in \mathcal{L}(X)$  with length  $\ell$  for which  $\mu([y]) = 0$ . Without repeating the details, we use topological transitivity to construct a map  $h$  as before, restricting the domain to a set of pairs which use the same gap as in the definition of  $g$  just above. Rather than (11), we get the following slightly weaker bound.

$$(14) \quad |\mathcal{L}_n(X)| \geq \left\lfloor \frac{n - 2f(n) - \ell}{f(n)(f(n) + \ell)} \right\rfloor e^{h(X)(n - 2f(n) - \ell)}.$$

However, (14) still contradicts Theorem 3.1 for large enough  $n$  (and small enough  $C$ ) since  $\frac{f(n)}{\ln n} \rightarrow 0$ . Therefore, our assumption was incorrect, and  $\mu$  is fully supported, completing the proof.  $\square$

**Remark 3.4.** The same techniques used in the proofs of Theorems 3.2 and Theorem 1.4 would allow a version of Theorem 1.3 assuming a weaker version of non-uniform two-sided specification, where one only assumes that any finite set of words can be combined for a single choice of gaps which are each less than or equal to the corresponding thresholds for each consecutive pair of words. This would cause an extra factor of  $f(n)$  in the upper bound from Theorem 3.1 and of  $(f(n))^3$  in the denominator in the right-hand side of (9), which do not affect the proof. We did not explicitly state this result because we did not want to add to the already large list of subtly different properties considered in this paper.

#### 4. EXAMPLES (PROOFS OF THEOREMS 1.5, 1.8, AND 1.9)

We begin with Theorem 1.5, yielding a class of subshifts to which Theorem 1.4 can be applied.

*Proof of Theorem 1.5.* Choose any such  $\{w_n\}$  and define  $X := X(\{w_n\})$ . We will prove topological mixing of  $X$  via the following statement.

**Claim 1:** For every  $n$ , every  $v, w \in \mathcal{L}_{|w_n|}(X)$ , and any  $\ell \in [2 \cdot 3^{n+1}, 2 \cdot 3^{n+2})$ , there exists  $u$  with  $|u| = \ell$  for which  $vwu \in \mathcal{L}(X)$ .

We begin with some auxiliary definitions. For any word  $w$ , we define functions  $p(w)$  and  $s(w)$  which measure how close  $w$  is to the creation of a forbidden word. Formally,  $p : \{0, 1\}^* \rightarrow \mathbb{N}^{\mathbb{N}}$  is defined as follows: for each  $n$ ,  $(p(w))(n)$  is the maximum length of a prefix of  $w$  which is also a subword of  $w_n$ . The function  $s$  is similarly defined; the only difference is that “prefix” is replaced by “suffix.” Clearly,  $0 \leq (p(w))(n), (s(w))(n) \leq \min(|w|, |w_n|)$  for each  $n$ . We note that  $(p(w))(n) = |w_n|$  iff  $w$  has  $w_n$  as a prefix, and  $(s(w))(n) = |w_n|$  iff  $w$  has  $w_n$  as a suffix.

We need some simple facts about the behavior of  $p, s$  when single letters are prepended or appended to a word; we begin with  $s$ . It should be clear that  $(s(wa))(n), (s(aw))(n) \leq (s(w))(n) + 1$  for all  $n, w$ , and  $a$ . We also claim that for every  $w$ , there exists  $a$  so that  $(s(wa))(n) \leq \frac{1}{3}|w_n|$ . If  $(s(w))(n) < \frac{1}{3}|w_n|$ , then this is true for any choice of  $a$  since  $(s(wa))(n) \leq (s(w))(n) + 1$ . If  $(s(w))(n) \geq \frac{1}{3}|w_n|$ , then  $w$  has a suffix  $s$  of length at least  $\frac{1}{3}|w_n|$  which is a subword of  $w_n$ . It cannot be the case that both  $s0$  and  $s1$  are subwords of  $w_n$ , since this would contradict

hypothesis (3) on  $w_n$ . Therefore, there exists  $a$  so that  $sa$  is not a subword of  $w_n$ , meaning that  $(s(wa))(n) \leq \frac{1}{3}|w_n|$ . For later reference, we say that any such  $a$  is a “right  $n$ -breaking” letter for  $w$ . Finally, we note that when  $|w| \geq |w_n|$ ,  $(s(aw))(n) = (s(w))(n)$ ; for such  $w$ , any subword of  $w_n$  is a suffix of  $w$  iff it is a suffix of  $aw$ . We omit the similar proofs that  $(p(aw))(n), (p(wa))(n) \leq (p(w))(n) + 1$ , that for every  $w$  there exists at least one “left  $n$ -breaking letter”  $a$  for which  $(p(aw))(n) \leq \frac{1}{3}|w_n|$ , and that when  $|w| \geq |w_n|$ ,  $(p(wa))(n) = (p(w))(n)$ .

The main step in the proof of Claim 1 is the following.

**Claim 2:** For any  $w \in \mathcal{L}_{|w_n|}(X)$  and for any  $\ell \in [3^{n+1}, 3^{n+2}]$ , there exist words  $t, u$  of length  $\ell$  so that  $tw, wu$  contain no word  $w_i$  for  $1 \leq i \leq n$ , and for which  $(p(tw))(i), (s(wu))(i) < \frac{1}{2}|w_i|$  for  $1 \leq i \leq n$ . (This implies as well that for  $1 \leq i \leq n$ ,  $twu$  contains no word  $w_i$  and  $(p(twu))(i), (s(twu))(i) < \frac{1}{2}|w_i|$ , since  $|w| = |w_n|$ .)

We prove Claim 2 by induction on  $n$ . In all cases, we describe only the proof of existence of  $u$ , as that of  $t$  is trivially similar. The base case  $n = 1$  is simple; for any  $w \in \mathcal{L}_{|w_1|}(X)$ , simply define  $u$  by following  $w$  by  $\ell$  right 1-breaking letters. Then  $(s(wu))(1) \leq \frac{1}{3}|w_1| < \frac{1}{2}|w_1|$ ,  $w$  contained no occurrence of  $w_1$  since  $w \in \mathcal{L}(X)$ , and none of the right 1-breaking letters could be the last letter of an occurrence of  $w_1$  by definition.

Now, for any  $n > 1$ , assume Claim 2 for  $n - 1$ , and choose  $\ell \in [3^{n+1}, 3^{n+2}]$ . We break into cases depending on whether  $(s(w))(n) < |w_n| - 3^n$  or  $(s(w))(n) \geq |w_n| - 3^n$ . First, suppose that  $(s(w))(n) < |w_n| - 3^n$ . Then, consider the suffix  $w'$  of  $w$  with length  $|w_{n-1}|$ . Clearly  $w' \in \mathcal{L}(X)$ , and so by the inductive hypothesis there exists  $u'$  with length  $3^n$  for which  $w'u'$  contains no word  $w_i$  for  $1 \leq i \leq n - 1$ , and  $(s(w'u'))(i) < \frac{1}{2}|w_i|$  for  $1 \leq i \leq n - 1$ . Since  $|w'| = |w_{n-1}|$  and  $w$  contains no  $w_i$  for  $1 \leq i \leq n - 1$  (it's in  $\mathcal{L}(X)$ ),  $wu'$  also contains no  $w_i$  for  $1 \leq i \leq n - 1$ . Since  $|w'u'| \geq |w_{n-1}|$ ,  $(s(wu'))(i) = (s(w'u'))(i) < \frac{1}{2}|w_i|$  for  $1 \leq i \leq n - 1$  as well. Finally, for any prefix  $p$  of  $u'$ ,  $(s(wp))(n) \leq (s(w))(n) + |p| < |w_n|$ , so  $wu'$  contains no  $w_n$ .

We now follow  $wu'$  by a word  $u''$  of length  $\ell - 3^n > 2^n$  made up of letters which are right  $j$ -breaking for  $j$  following the pattern  $n, 1, 2, 1, 3, 1, 2, 1, 4, \dots$ . By this, we mean that each letter is right  $j$ -breaking for the proper  $j$  and for the word made up of  $wu'$  followed by the previously placed letters. We claim that  $wu'u''$  contains no word  $w_i$  for  $1 \leq i \leq n$ , and that  $(s(wu'u''))(i) < \frac{1}{2}|w_i|$  for  $1 \leq i \leq n$ . Since  $wu'$  contains no  $w_i$  for  $1 \leq i \leq n$ , if  $wu'u''$  did contain some  $w_i$ , there would exist a prefix  $p$  of  $u''$  for which  $wu'p$  has  $w_i$  as a suffix. However, the final letter of  $p$  has distance at most  $2^i$  from either the end of  $wu'$  or from the nearest occurrence of a right  $i$ -breaking letter to the left. Therefore  $(s(wu'p))(i) \leq \max((s(wu'))(i), \frac{1}{3}|w_i|) + 2^i \leq \frac{1}{2}|w_i| + 2^i < |w_i|$  by hypothesis (2), a contradiction to  $wu'p$  ending with  $w_i$ . In particular, for each  $1 \leq i \leq n$ , the rightmost letter in  $u''$  has distance less than or equal to  $2^i$  from some  $i$ -breaking letter, and so  $(s(wu'u''))(i) \leq \frac{1}{3}|w_i| + 2^i < \frac{1}{2}|w_i|$  for each  $i$  by hypothesis (2). We can then define  $u = u'u''$  to satisfy the conclusion of Claim 2.

If instead  $(s(w))(n) \geq |w_n| - 3^n$ , then we use the fact that  $w \in \mathcal{L}(X)$  to find a  $3^n$ -letter word  $x$  for which  $wx \in \mathcal{L}(X)$ . Some letter of  $wx$  must be right  $n$ -breaking for the portion of  $wx$  preceding it, else for some prefix  $p$  of  $x$  we have

$(s(wp))(n) = |w_n|$ , meaning that  $wp$  (and so also  $wx$ ) contains  $w_n$ , a contradiction to  $wx \in \mathcal{L}(X)$ . Define  $x'$  to be a prefix of  $x$  ending with such a right  $n$ -breaking letter; then  $(s(wx'))(n) \leq \frac{1}{3}|w_n|$ . Then, as above, we apply the inductive hypothesis to the suffix of length  $|w_{n-1}|$  of  $wx'$  to find  $x''$  with length  $3^n$  so that  $wx'x''$  contains no  $w_i$  for  $1 \leq i \leq n$  and  $(s(wx'x''))(i) < \frac{1}{2}|w_i|$  for  $1 \leq i \leq n-1$ . Note that also  $(s(wx'x''))(n) \leq (s(wx'))(n) + 3^n \leq \frac{1}{3}|w_n| + 3^n < \frac{1}{2}|w_n|$  by hypothesis (2). Now, we proceed exactly as in the previous case, following by a word  $x'''$  of length  $\ell - 2 \cdot 3^n > 2^n$  made of right  $i$ -breaking letters for  $i$  following the pattern  $n, 1, 2, 1, 3, 1, 2, 1, 4, \dots$ . Exactly as in the previous case,  $wx'x''x'''$  contains no word  $w_i$ , and  $(s(wx'x''x'''))(i) < \frac{1}{2}|w_i|$  for  $1 \leq i \leq n$ . We can then take  $u = x'x''x'''$  as in the conclusion of Claim 2. In each case, we've found  $u$  as in Claim 2, completing the proof of Claim 2 by induction.

We now prove Claim 1. Take any  $v, w \in \mathcal{L}_{|w_n|}(X)$  and any  $\ell \in [2 \cdot 3^{n+1}, 2 \cdot 3^{n+2}]$ ; partition  $\ell = \ell' + \ell''$  for  $\ell', \ell'' \in [3^{n+1}, 3^{n+2}]$ . By Claim 2, there exist  $t', u'$  of length  $\ell'$  and  $t'', u''$  of length  $\ell''$  so that  $t'vu', t''wu''$  contain no  $w_i$  for  $1 \leq i \leq n$ , and  $(p(t'vu'))(i), (s(t'vu'))(i), (p(t''wu''))(i), (s(t''wu''))(i) < \frac{1}{2}|w_i|$  for  $1 \leq i \leq n$ . We now consider the word  $t'vu't''wu''$ . First, we claim that it contains no  $w_i$  for  $1 \leq i \leq n$ . To see this, note that if it did contain some  $w_i$ , it could not be contained in either  $t'vu'$  or  $t''wu''$ , and then either  $t'vu'$  ends with a subword of  $w_i$  of length at least  $\frac{1}{2}|w_i|$  (contradicting  $(s(t'vu'))(i) < \frac{1}{2}|w_i|$ ) or  $t''wu''$  begins with a subword of  $w_i$  of length at least  $\frac{1}{2}|w_i|$  (contradicting  $(p(t''wu''))(i) < \frac{1}{2}|w_i|$ ). We also note that  $|t'vu't''wu''| \leq 2|w_n| + 4 \cdot 3^{n+2} < |w_{n+1}|$  by hypothesis (2), and so  $(p(t'vu't''wu''))(i), (s(t'vu't''wu''))(i) < \frac{1}{2}|w_i|$  for  $1 \leq i \leq n$ . In fact, since  $2|w_n| + 4 \cdot 3^{n+2} < \frac{1}{2}|w_{n+1}|$  by hypothesis (2),  $(p(t'vu't''wu''))(i), (s(t'vu't''wu''))(i) < \frac{1}{2}|w_i|$  for all  $i$ . We are now ready to show that  $t'vu't''wu'' \in \mathcal{L}(X)$ , which will complete our proof since it would imply that  $vu'tu''w \in \mathcal{L}(X)$  and  $|u't''| = \ell$ .

The idea is to add left and right breaking letters similarly to the proof of Claim 2. We alternate between placing letters to the right of  $t'vu't''wu''$  which are right  $j$ -breaking for  $j$  following the pattern  $1, 2, 1, 3, 1, 2, 1, 4, \dots$  and to the left which are left  $j$ -breaking for  $j$  following the reversed pattern  $\dots, 4, 1, 2, 1, 3, 1, 2, 1$ . For each placed letter, when we say it is left or right  $j$ -breaking, we mean that it is left or right  $j$ -breaking for the word comprised of  $t'vu't''wu''$  together with all previously placed letters to its left and right. This process yields a point  $x \in \{0, 1\}^{\mathbb{Z}}$ . We claim that  $x \in X$ , which will imply that  $vu'tu''w \in \mathcal{L}(X)$ .

For this, we need only show that  $x$  contains no  $w_i$ . Suppose for a contradiction that  $x$  contains some  $w_i$ . Define by  $z$  the first word during the construction of  $x$  containing  $w_i$ . Then  $z$  has  $w_i$  as either a suffix or prefix; we treat only the former case as the latter is trivially similar. By construction, either  $z$  occurs within the first  $2^{i+1}$  steps of the construction of  $x$ , or within  $2^{i+1}$  steps of the placement of a right  $i$ -breaking letter. Therefore,  $z = qyr$ , where  $|q| + |r| \leq 2^{i+1}$  and  $(s(y))(i) < \frac{1}{2}|w_i|$  (either because  $y = t'vu't''wu''$  or because  $y$  ended with a right  $i$ -breaking letter). Then,  $(s(z))(i) \leq \frac{1}{2}|w_i| + 2^{i+1} < |w_i|$  by hypothesis (2), a contradiction to  $z$  having  $w_i$  as a suffix. We've shown that  $x$  contains no  $w_i$ , so  $x \in X$  and  $vu'tu''w \in \mathcal{L}(X)$ , completing the proof of Claim 1.

It remains only to show that Claim 1 implies the desired topological mixing of  $X$ . Consider any words  $v, w \in \mathcal{L}(X)$  with  $|v|, |w| \leq |w_n|$  and any  $\ell > 2 \cdot 3^{n+1}$ .



Choose  $n'$  so that  $\ell \in [2 \cdot 3^{n'+1}, 2 \cdot 3^{n'+2}]$ ; clearly  $n' \geq n$ . Then, there exist  $v', w' \in \mathcal{L}_{|w_n|}(X)$  where  $v'$  ends in  $v$  and  $w'$  starts with  $w$ . By Claim 1, there exists  $u$  with  $|u| = \ell$  for which  $v'uw' \in \mathcal{L}(X)$ , and then clearly  $vu w \in \mathcal{L}(X)$  as well. This proves that  $X$  is topologically mixing with gap function  $f$  defined by  $f(n) = 2 \cdot 3^{k+1}$  for  $n \in (|w_{k-1}|, |w_k|]$ . It is clear that  $\lim_{n \rightarrow \infty} \frac{f(n)}{\ln n} = 0$  given hypothesis (1) on  $|w_n|$ , completing the proof of the theorem.  $\square$

Next, we prove Theorem 1.8, demonstrating that topological mixing with slow growth along a subsequence cannot in general imply intrinsic ergodicity (or even positive topological entropy!).

*Proof of Theorem 1.8.* Choose any increasing unbounded function  $f(n)$ . Define increasing sequences of integers  $(m_k)$  and  $(n_k)$  as follows. First, define  $n_1 = m_1 = 1$ . Then, for each  $k > 1$ , choose  $m_k$  to be greater than  $n_{k-1} + 3m_{k-1}$  and large enough that  $f(m_k) > kn_{k-1}$ , and define  $n_k = km_k$ . Define a subshift  $X$  on  $\{0, 1\}$  as follows: for  $x \in \{0, 1\}^{\mathbb{N}}$ ,  $x$  is in  $X$  iff for every  $k$ , every  $(n_k + 3m_k)$ -letter subword of  $x$  contains at least  $n_k$  letters which are equal to their neighbor to the right.

We claim that the only invariant measures on  $X$  are convex combinations of the  $\delta$ -measures on the points  $0^{\mathbb{Z}}, 1^{\mathbb{Z}}$ . To see this, suppose that  $\mu$  is an invariant measure on  $X$ . Then, for every  $x \in X$ , the ergodic averages

$$\frac{1}{n_k + 3n_k} \sum_{i=0}^{n_k + 3m_k - 1} \chi_{[00] \cup [11]}(\sigma^i x)$$

converge to 1 as  $k \rightarrow \infty$  since  $\frac{m_k}{n_k} \rightarrow 0$ . Therefore, by Birkhoff's ergodic theorem,  $\mu([00] \cup [11]) = 1$ , and so  $\mu$ -a.e.  $x \in X$  has all digits equal to their neighbor on the right. This means that  $\mu$  is a convex combination of  $\delta_{0^{\mathbb{Z}}}$  and  $\delta_{1^{\mathbb{Z}}}$  as claimed. Since all such measures have entropy 0, by the variational principle,  $h(X) = 0$ , and so  $\delta_{0^{\mathbb{Z}}}$  and  $\delta_{1^{\mathbb{Z}}}$  are ergodic measures of maximal entropy.

We now claim that for every  $k$ , any words  $v, w \in \mathcal{L}_{m_k}(X)$ , and any  $\ell \geq 2n_{k-1}$ , there exists  $u \in \mathcal{L}_{\ell}(X)$  so that  $vu w \in \mathcal{L}(X)$ . To see this, choose any  $k$  and any  $v, w \in \mathcal{L}_{m_k}(X)$ . Define the maximal constant prefix and suffix  $p_v, s_v$  of  $v$  and the maximal constant prefix and suffix  $p_w, s_w$  of  $w$ . Say that the constant words  $p_v, s_v, p_w, s_w$  contain the letters  $\pi_v, \sigma_v, \pi_w, \sigma_w$  respectively. We now claim that for every  $\ell \geq 2n_{k-1}$ ,  $\pi_v^{\infty} v \sigma_v^{n_{k-1}-1} \pi_w^{\ell-n_{k-1}} w \sigma_w^{\infty} \in X$ , which will imply that  $v \sigma_v^{n_{k-1}-1} \pi_w^{\ell-n_{k-1}} w \in \mathcal{L}(X)$ , allowing us to take  $u = \sigma_v^{n_{k-1}-1} \pi_w^{\ell-n_{k-1}}$ .

For this proof, we need to show that for every  $j$  and every  $(n_j + 3m_j)$ -letter subword  $z$  of  $\pi_v^{\infty} v \sigma_v^{n_{k-1}-1} \pi_w^{\ell-n_{k-1}} w \sigma_w^{\infty}$ ,  $z$  contains at least  $n_j$  letters equal to their neighbor on the right. We break into cases, and first treat the case where  $j \geq k$ . Then  $z$  has length  $n_j + 3m_j > n_j + 2m_k$ . However, the only letters of  $\pi_v^{\infty} v \sigma_v^{n_{k-1}-1} \pi_w^{\ell-n_{k-1}} w \sigma_w^{\infty}$  which are not equal to their neighbor on the right are either the final letter of  $\sigma_v^{n_{k-1}-1}$  or non-terminal letters of  $v$  or  $w$ , and the total number of such letters is less than  $2m_k$ . So,  $z$  has at least  $n_j$  letters equal to their neighbors on the right.

Now assume that  $j < k$ , and again  $|z| = n_j + 3m_j \leq n_{k-1} + 3m_{k-1} < m_k$ . If  $z$  is a subword of either  $v$  or  $w$ , then  $z$  contains at least  $n_j$  letters equal to their neighbor on the right since  $v, w \in \mathcal{L}(X)$ . If  $z$  contains either  $\sigma_v^{n_{k-1}-1}$  or  $\pi_w^{\ell-n_{k-1}}$ , then  $z$  contains  $n_{k-1}$  letters which are equal to their neighbors on the right (recall that  $v$  ends with  $\sigma_v$  and  $w$  begins with  $\pi_w$ ). If  $z$  is contained in  $\sigma_v^{n_{k-1}-1} \pi_w^{\ell-n_{k-1}}$ , then  $z$  has

at least  $n_j$  letters equal to their neighbor on the right; in fact it has at most one letter without this property. The only remaining case is that  $z = qr$ , where either  $r$  is a prefix of  $v$  and  $q$  has all letters  $\pi_v$ ,  $q$  is a suffix of  $v$  and  $r$  has all letters  $\sigma_v$ ,  $r$  is a prefix of  $w$  and  $q$  has all letters  $\pi_w$ , or  $q$  is a suffix of  $w$  and  $r$  has all letters  $\sigma_w$ . We treat only the first case, as the others are extremely similar. Note that since  $r \in \mathcal{L}(X)$ , there exists  $q'$  with length  $|q|$  so that  $q'r \in \mathcal{L}(X)$ ; in particular,  $q'r$  contains at least  $n_j$  letters equal to their neighbor on the right. Obviously  $qr$  contains at least as many letters equal to their neighbors on the right as  $q'r$ , and so  $qr$  has at least  $n_j$  letters equal to their neighbor on the right.

We have shown that  $\pi_v^\infty v \sigma_v^{n_{k-1}} \pi_w^{\ell-n_{k-1}} w \sigma_w^\infty \in X$ , and so verified the claim that for every  $k$ , any words  $v, w \in \mathcal{L}_{m_k}(X)$ , and any  $\ell \geq 2n_{k-1}$ , there exists  $u \in \mathcal{L}_\ell(X)$  so that  $vuw \in \mathcal{L}(X)$ . Clearly then, for any words  $v, w \in \mathcal{L}_n(X)$  with  $n \leq m_k$  and for any  $\ell \geq 2n_{k-1}$ , we can extend  $v$  on the left and  $w$  on the right to words in  $\mathcal{L}_{m_k}(X)$ , and we will again find  $u \in \mathcal{L}_\ell(X)$  for which  $vuw \in \mathcal{L}(X)$ . Therefore  $X$  is topologically mixing for the function  $g$  defined by  $g(n) = 2n_{k-1}$  for  $n \in (m_{k-1}, m_k]$ . Since we assumed that  $f(m_k) > kn_{k-1}$ ,  $\frac{g(m_k)}{f(m_k)} = \frac{2n_{k-1}}{f(m_k)} \rightarrow 0$  and so  $\liminf \frac{g(n)}{f(n)} = 0$ .  $\square$

Finally, we prove Theorem 1.9, showing that for unbounded gap functions, topological mixing never implies the existence of periodic points.

*Proof of Theorem 1.9.* We construct a class of examples with no periodic points, and show that the class contains subshifts with topological mixing for arbitrarily slowly growing gap function. Choose any sequence  $\{n_k\}$  with the property that  $n_k > 2n_{k-1} + 2k$  for all  $k$ .

Fix any subshift  $S$  on the alphabet  $\{0, 1\}$  without periodic points. We define  $X$  on the same alphabet  $\{0, 1\}$  by the following rule: a point  $x \in \{0, 1\}^{\mathbb{Z}}$  is in  $X$  if and only if for every  $k$ , every  $(n_k + 2k)$ -letter subword of  $x$  contains a subword in  $\mathcal{L}_k(S)$ . Then clearly  $X$  has no periodic points, since every point of  $X$  contains arbitrarily long words in  $\mathcal{L}(S)$ . We claim that in addition,  $X$  is topologically mixing with gap function depending on  $\{n_k\}$ .

To show this, consider any  $k$ , any two words  $v, w \in \mathcal{L}_{n_k}(X)$ , and any  $\ell \geq 2(k-1)$ ; we will show that there exists  $u$  with length  $\ell$  so that  $vuw \in \mathcal{L}(X)$ . Assume first that  $\ell$  is even.

Of the prefixes of  $v$ , choose the one with maximal length which is in  $\mathcal{L}(S)$ , and denote it by  $p_v$ . Similarly define a prefix  $p_w$  of  $w$ , and suffixes  $s_v$  and  $s_w$  of  $v$  and  $w$  respectively. Since  $p_v, s_w \in \mathcal{L}(S)$ , there exist a left-infinite sequence  $x$  and a right-infinite sequence  $y$  for which  $xp_v, s_wy \in \mathcal{L}(S)$ . Similarly, since  $s_v, p_w \in \mathcal{L}(S)$ , there exist  $s, t$  with length  $\ell/2 \geq k-1$  so that  $s_v s, t p_w \in \mathcal{L}(S)$ . We claim that  $xvstwy \in X$ , which will imply that  $vstwy \in \mathcal{L}(X)$ , completing our proof by taking  $u = st$ .

For this proof, we need to show that for every  $j$  and every  $(n_j + 2j)$ -letter subword  $z$  of  $xvstwy$ ,  $z$  contains a word in  $\mathcal{L}_j(S)$ . We break into cases, and first treat the case where  $j > k$ . Then  $z$  has length  $n_j + 2j > 2n_{j-1} + 4j \geq 2n_k + 4j$ . If  $\ell/2 \geq j$ , then  $z$  must contain a  $j$ -letter subword of one of  $s, t, x$ , or  $y$ , and then it contains a word in  $\mathcal{L}_j(S)$  by definitions of  $s, t, x$ , and  $y$ . If  $\ell/2 < j$ , then  $z$  must contain a  $j$ -letter subword of either  $x$  or  $y$ , and again it then contains a word in  $\mathcal{L}_j(S)$ .

Suppose instead that  $|z| = n_j + 2j$  for  $j \leq k$ . Clearly, if  $z$  is a subword of either  $v$  or  $w$ , then  $z$  contains a word in  $\mathcal{L}_j(S)$  since  $v, w \in \mathcal{L}(X)$ . Similarly, if  $z$  contains

a  $j$ -letter subword of  $x$ ,  $y$ ,  $s_vs$ , or  $tp_w$ , then again it contains a word in  $\mathcal{L}_j(S)$ . If  $z$  contains letters from both  $s$  and  $t$ , then  $z$  contains a  $j$ -letter subword of either  $s$  or  $t$ , which in either case is a word in  $\mathcal{L}_j(S)$ . The only remaining case is that  $z = qr$ , where either  $|q| < j$  and  $r$  is a prefix of  $v$ ,  $|q| < j$  and  $r$  is a prefix of  $w$ ,  $q$  is a suffix of  $v$  and  $|r| < j$ , or  $q$  is a suffix of  $w$  and  $|r| < j$ . Note that this cannot happen for  $j = k$ , and so now we know that  $j < k$ . We treat only the case  $|q| < j$  and  $r$  a prefix of  $v$ , as the others are extremely similar. Recall that  $p_v$  is the prefix of  $v$  of maximal length which is in  $\mathcal{L}(S)$ .

If  $|q| \geq j - |p_v|$ , then  $z = qr$  contains a  $j$ -letter subword of  $xp_v \in \mathcal{L}(S)$ , which is in  $\mathcal{L}_j(S)$ . The only remaining case is that  $|q| < j - |p_v|$ , implying that  $|p_v| < j$  and that  $|r| > n_j + j + |p_v|$ . We claim that in this case,  $r$  actually contains a word in  $\mathcal{L}_j(S)$ , which obviously implies that  $z$  does as well. It suffices to show that the  $(n_j + j + |p_v| + 1)$ -letter prefix  $r'$  of  $v$  contains a word in  $\mathcal{L}_j(S)$ , since  $r$  contains  $r'$ . Recall that  $v$  was in  $\mathcal{L}(X)$ , so there exists  $q'$  with length  $j - |p_v| - 1$  for which  $q'r' \in \mathcal{L}(X)$ . In particular, this means that  $q'r'$  contains a word in  $\mathcal{L}_j(S)$ . However, this word in  $\mathcal{L}_m(S)$  cannot contain any letters of  $q'$ ; if it did, then it would by necessity contain the prefix of  $r'$  of length  $|p_v| + 1$ , which by definition of  $p_v$  was not in  $\mathcal{L}(S)$ , a contradiction. Therefore,  $r'$  contains that word in  $\mathcal{L}_j(S)$ , and we are done. As stated earlier, the other possibilities for  $q, r$  are similar, and so we have shown that  $xvstwy \in X$ , and in particular that for even  $\ell \geq 2k$ , that there exists  $u$  with length  $2(k - 1)$  so that  $vuuv \in \mathcal{L}(X)$ . The proof for odd  $\ell \geq 2(k - 1)$  requires nearly no change; one just takes the lengths of  $s$  and  $t$  to be  $\lfloor \ell/2 \rfloor$  and  $\lceil \ell/2 \rceil$  respectively.

Since any  $u, v \in \mathcal{L}(X)$  with lengths less than or equal to  $n_k$  can be extended on the left and right to words in  $\mathcal{L}_{n_k}(X)$ ,  $X$  is topologically mixing with gap function  $g(n)$  which is equal to 0 for  $0 < n \leq n_1$  and  $2k$  for  $n_k < n \leq n_{k+1}$ . We now note that this  $g$  can be chosen less than or equal to an arbitrary unbounded function  $f$ , for instance by defining  $\{n_k\}$  so that for all  $k$ ,  $f(n_k) > 2k$ . Then, since topological mixing with gap function  $g$  implies topological mixing for  $f \geq g$ , we are done.  $\square$

## REFERENCES

- [1] Rufus Bowen. Some systems with unique equilibrium states. *Math. Systems Theory*, 8(3):193–202, 1974/75.
- [2] Vaughn Climenhaga and Daniel J. Thompson. Intrinsic ergodicity beyond specification:  $\beta$ -shifts,  $S$ -gap shifts, and their factors. *Israel J. Math.*, 192(2):785–817, 2012.
- [3] Masahito Dateyama. The almost weak specification property for ergodic group automorphisms of abelian groups. *J. Math. Soc. Japan*, 42(2):341–351, 1990.
- [4] Manfred Denker, Christian Grillenberger, and Karl Sigmund. *Ergodic theory on compact spaces*. Lecture Notes in Mathematics, Vol. 527. Springer-Verlag, Berlin-New York, 1976.
- [5] B. M. Gurevich. Stationary random sequences of maximal entropy. In *Multicomponent random systems*, volume 6 of *Adv. Probab. Related Topics*, pages 327–380. Dekker, New York, 1980.
- [6] Nicolai T. A. Haydn. Phase transitions in one-dimensional subshifts. *Discrete Contin. Dyn. Syst.*, 33(5):1965–1973, 2013.
- [7] Wolfgang Krieger. On the uniqueness of the equilibrium state. *Math. Systems Theory*, 8(2):97–104, 1974/75.
- [8] Dominik Kwietniak, Piotr Oprocha, and Michał Rams. On entropy of dynamical systems with almost specification. *Israel J. Math.*, 213(1):475–503, 2016.
- [9] Douglas Lind and Brian Marcus. *An introduction to symbolic dynamics and coding*. Cambridge University Press, Cambridge, 1995.

- [10] Brian Marcus. A note on periodic points for ergodic toral automorphisms. *Monatsh. Math.*, 89(2):121–129, 1980.
- [11] Joseph S. Miller. Two notes on subshifts. *Proc. Amer. Math. Soc.*, 140(5):1617–1622, 2012.
- [12] Ronnie Pavlov. On intrinsic ergodicity and weakenings of the specification property. *Adv. Math.*, 295:250–270, 2016.
- [13] Peter Walters. *An Introduction to Ergodic Theory*. Number 79 in Graduate Texts in Mathematics. Springer-Verlag, 1982.
- [14] Kôzaku Yosida and Shizuo Kakutani. Birkhoff's ergodic theorem and the maximal ergodic theorem. *Proc. Imp. Acad., Tokyo*, 15:165–168, 1939.

RONNIE PAVLOV, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF DENVER, 2280 S. VINE ST., DENVER, CO 80208

*E-mail address:* `rpavlov@du.edu`

*URL:* `www.math.du.edu/~rpavlov/`